Atlantis Highlights in Engineering (AHE), volume 2
International Conference on Industrial Enterprise and System Engineering (IcoIESE 2018)

# Copyright Protection High Robustness Arnold Transform Wavelet Method

Bruguiera A. F. Agradriya[1], Irma Safitri [2], Ledya Novamizanti [3] and Muhammad Ainul Yaqin [4]
[1,2,3,4]School of Electrical Engineering, Telkom University
Bandung, Indonesia
{bruguieraafa[1], ledyamizan[3], yaqinmuhammad.22[4]}@gmail.com, irmasaf@telkomuniversity.ac.id[2]

***Abstract*— Audio watermarking is the technique of information insertion into an audio signal without altering the significant origin of the audio host. Audio watermarking is usually used to protect the copyright of a work in the form of songs, recording of state and other secrets to avoid irresponsible persons. Audio watermarking technique is passed through two stages including embedding, and extracting. The extracted audio signal must be resistant to several attacks. This research analyzes the third scheme of audio watermarking system on Arnold Transform with Discrete Wavelet Transform-Singular Value Decomposition (DWT-SVD) methods. The results show that the systems we designed are robust to some attack types, namely LPF 10 kHz, cropping, MP3 compression 256 kHz, and MP3 compression 512 kHz. Our system has BER=0.0012, PSNR=29.2082 and CC=0.9976 for compression and LPF attacks; while with cropping attack this paper shows BER=0, PSNR=∞ and CC=1.***

***Keywords*— Audio Watermarking; Arnold Transform; Discrete Wavelet Transform; Singular Value Decomposition.**

## I. INTRODUCTION

Technological developments have greatly facilitated the ease of access and manipulation of multimedia data such as text, images, audio or video. This has caused serious problems for copyright protection. This issue is so crucial that it takes technology to protect digital content to prevent copyright infringement.

Digital watermarking is one solution to prevent copyright infringement where the information signal inserted into the host signal can be text, image, audio or video. This insertion is done in such a way that the inserted information does not interfere with the use of the data host [1]. In this paper we will discuss audio watermarking. Audio watermarking is using techniques by utilizing the weakness of the human auditory sense that is HAS (Human Auditory System).

Audio watermarking can operate on two domains, the frequency domain and time domain. Each domain has its own characteristics. The frequency domain is more effective than the time domain because at the time of the watermarking process into the audio signal, the inaudibility and robustness factor can be maintained [2]. An audio watermarking can be considered feasible if some of the following factors must be met namely inaudibility, robustness and security. Inaudibility is a condition in which the embedded watermark should not issue distortion to the original audio signal, whereas robustness determines whether the watermark signal is resistant to degradation or removal [3].

Audio watermarking has several methods in the previous research. These techniques are such as DCT [4, 5], DWT [5-9, 11-13], the combination of DWT-SVD [5, 7, 9, 11, 13], etc. In [13] we proposed 2 schemes of Arnold transform audio watermarking, but it does not include the attack atached to the audio watermarking to analyze the robustness. In this work, we extend the previous work and analyze the robustness of our system.

## II. AUDIO WATERMARKING METHOD

### A. Singular Value Decomposition (SVD)

SVD is a numerical analysis tool that uses decomposition of semi-definite-positive matrix eigenvalues to obtain the same decomposition [10]. It can be used for all real numbered rectangular matrices. Thus, if SVD is done in a semi-definite-positive matrix it will give the same result as eigenvalue decomposition. SVD has an important character in providing optimum estimates from a matrix through a matrix that has a smaller rank. SVD can provide the best estimate of a quadrilateral matrix with other quadrilateral matrix. If there is a quadrilateral matrix A, SVD of the A matrix can be expressed as in equation (1):

$$
\begin{aligned}
A &= USV^T \qquad (1)\\
&= \begin{bmatrix} u_{1,1} & \cdots & u_{1,r} \\ \vdots & \ddots & \vdots \\ u_{m,1} & \cdots & u_{m,r} \end{bmatrix} \times \begin{bmatrix} s_{1,1} & \cdots & s_{1,r} \\ \vdots & \ddots & \vdots \\ s_{m,1} & \cdots & s_{m,r} \end{bmatrix} \times \begin{bmatrix} v_{1,1} & \cdots & v_{1,r} \\ \vdots & \ddots & \vdots \\ v_{m,1} & \cdots & v_{m,r} \end{bmatrix}^T \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{r} u_{i,k} \times s_{k,k} \times v_{k,j}
\end{aligned}
$$

U is a matrix of mxr, V is a matrix of rxn and S is a diagonal matrix of rxr, and T is a transpose matrix.

Some characteristics of SVD are:

- When the singular value of change is small, the change of the large singular value is not included in the main matrix after inverse transformation
- Changes from a large singular value do not gain a place in any kind of audio signal processing.

### B. Arnold Transform

Arnold transform is widely used in digital watermarking technology [14]. It is a transformation based on Arnold scrambling (image scrambling) that has simple and periodic features. It means that after several cycles the properties of the inserted image will return to their original state. Arnold transform function of N×N dimensionless image is written in equation (2):

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} (\mathrm{mod}\, N_1) \quad (2)$$

and its inverse function is written as follows:

$$\begin{pmatrix} a_1' \\ b_1' \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} (\mathrm{mod}\, N_1) \quad (3)$$

### C. Discrete Wavelet Transform (DWT)

DWT represents a signal in several frequency components. The process obtaining these components is done by passing signals on the filter. The decomposition of signal is aimed to obtain signal composition at high and low frequencies. Wavelet decomposition passes the original audio signal through two filters, namely High Pass Filter (HPF) and Low pass Filter (LPF). This two filters output is down sampled to obtain detailed coefficient (D) for HPF and approximate coefficient (A) for LPF [3].

The component of the low frequency signal is focused on the energy of the audio signal, which is an important part of the audio signal, while the high frequency component only focuses on a fraction of the signal energy. Audio signals can be flexible and can be decomposed into multi-level DWT.

## III. SISTEM MODEL

We analyzed the third scheme of audio watermarking based on DWT-SVD-Arnorld Transform proposed in [13]. The watermark processes compose of embedding and extracting processes.

### A. Embedding Algorithm

Watermark embedding process is the process of inserting the watermark to the audio host by observing the intensity and frame values. In the embedding process we will change the intensity value, check the maximum bit capacity and maximum frame capacity. The detail processes can be seen in Figure 1.

### B. Extracting Algorithm

The process of watermark extraction is done to restore the audio host and the watermark signal. In this process we will know the robustness of the audio host and the watermark signal. The stages of complete extraction process can be seen in Figure 2.

## IV. ANALYSIS AND EXPERIMENT RESULTS

This paper applies the third scheme of the previously two schemes in [13]. In [13], the first scheme is DWT SVD audio watermarking without Arnold transform and the second one uses Arnold transform. In this paper, the results of watermarking will be inserted several attacks. This is done to test the robustness of the watermarking system against the

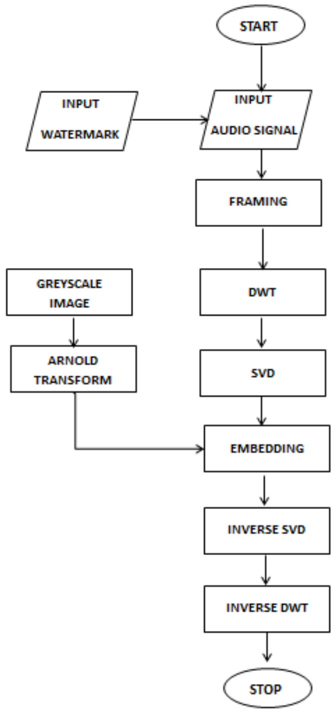



Fig. 1 Flow chart of embedding algorithm

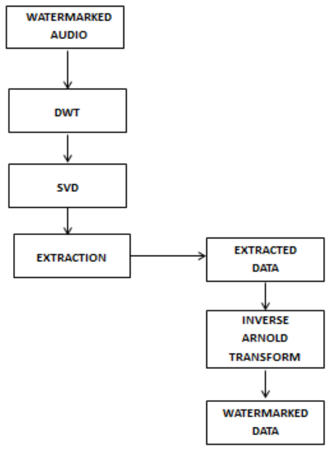



Fig. 2 Flow chart of extracting algorithm

attack. The attacks provided are LPF, stereo to mono, resampling, MP3 compression, noise and cropping.

Embedded signal used in this research is a 50x50 pixel gray scale image, while the audio host is in duration of 10 seconds with a frequency sampling of 44.1 KHz. We used 4 types of audio hosts: host2.wav, host3.wav, host4.wav, and host5.wav. Matlab is used in this reseach to simulate our audio watermarking system. Some parameters used in this paper are peak signal to noise ratio (PSNR) in equation (4) and correlation coefficient (CC) in equation (5).

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right) \quad (4)$$

$$\gamma = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2 \left(B_{mn} - \overline{B}\right)^2}} \quad (5)$$

4.1 Low Pass Filter (LPF)

The embedded audio file is given filtering attack, in this case LPF, with cut-off frequency used is 8 KHz, 9 KHz and 10 KHz. Table 1 shows the results of attacks on the parameters BER. Based on Table 1, file host4.wav at 8 KHz - 10 KHz cut-off frequency has BER value that is zero. This is because the embedded information is in a low sub-band so that when given a LPF attack this method has a satisfactory BER. Our system is robust to the LPF attack for host4.wav and host5.wav.

4.2 High Pass Filter (HPF)

In general, HPF is opposite to LPF, if LPF passes below the cut-off frequency while HPF passes above the cut-off frequency. Cut-off frequencies are 7 KHz, 12 KHz and 15 KHz. Table 2 shows the results of the effect of HPF attacks against the value of BER. In Table 2, all the tested files do not have BER = 0. This is due to the insertion of information performed at low sub-band frequencies. With the HPF properties passing above the cut-off frequency while the majority of information is below the cut-off frequency, it causes the value of BER to be non-zero. Our system is not robust against the HPF attack

4.3 Band Pass Filter (BPF)

The type of BPF attack passes between the lower cut-off frequency and upper cut-off frequency. The frequency ranges used in this paper are 100-2000 Hz, 100-4000 Hz, 100-6000 Hz and 100-8000 Hz. Table 3 shows the influence of BPF attacks on the BER value.

In Table 3, above all the tested files there is not a zero BER. However, the larger the frequencies range the BER decreases. This is due to the large cut-off frequency range that can escape the amount of information available in that range. This test system is not robust to BPF attacks.

4.4 Resampling

Resampling is a change in the number of signal samples in a watermarked audio. Generally resampling is divided into two, namely down sampling and up sampling. In testing this attack, our system has resampling attack ranging from the frequency of 8000 Hz, 11025 Hz, 22050 Hz and 88200 Hz. The results of testing the system against resampling attacks can be seen in Table 4. From the table, we have no zero BER value in each test file. It can be concluded that the watermarking audio system is less robust to resampling attacks.

4.5 MP3 Compression

MP3 compression is a compression type that works to reduce the size of the audio and produce MP3 extension files with a certain compression rate. In testing this attack, we change the bit rate of 128 Kbps, 256 Kbps, 320 Kbps and 512 Kbps. This analysis is shown in Table 5. According to Table 5, generally the test file shows a BER value of zero at the bit rate of 320 Kbps and 512 Kbps for all audio files. It can be concluded that the greater the bit rate the better the value of BER but the greater the size of the compression results. Our system is robust to the MP3 compression attack.

4.6 Noise

In the noise attack test, the watermarked audio is given a noise attack by changing the value of noise in decibels: -10 dB, -20 dB and -30 dB. The results can be seen in Table 6. The BER=0 can be achieved when the noise is -30 dB, which is robust in this noise level.

TABLE 1
EFFECT OF LPF ATTACKS ON BER VALUES

| Frequency cut-off | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 8 kHz | 0.0016 | 0.0116 | 0 | 0 |
| 9 kHz | 0.0016 | 0.0068 | 0 | 0 |
| 10 kHz | 0.0012 | 0.0048 | 0 | 0 |

TABLE 2
EFFECT OF HPF ATTACKS ON BER VALUES

| Frequency cut-off | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 7 KHz | 0.5112 | 0.5520 | 0.5556 | 0.5272 |
| 12 KHz | 0.5372 | 0.5564 | 0.5556 | 0.5556 |
| 15 KHz | 0.5504 | 0.5552 | 0.5556 | 0.5556 |

TABLE 3
EFFECT OF BPF ATTACKS ON BER VALUES

| Frequency cut-off | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 100-2000 Hz | 0.4932 | 0.5424 | 0.4848 | 0.5124 |
| 100-4000 Hz | 0.5056 | 0.1388 | 0.2556 | 0.3248 |
| 100-6000 Hz | 0.4684 | 0.1392 | 0.2396 | 0.2893 |
| 100-8000 Hz | 0.4772 | 0.1364 | 0.2364 | 0.2828 |

TABLE 4
EFFECT OF RESAMPLING ATTACKS ON BER VALUES

| Resampling | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 8000 Hz | 0.5095 | 0.5432 | 0.5482 | 0.5124 |
| 11025 Hz | 0.5089 | 0.5556 | 0.5210 | 0.5043 |
| 22050 Hz | 0.5000 | 0.5288 | 0.5155 | 0.5022 |
| 88200 Hz | 0.4912 | 0.5050 | 0.4916 | 0.4977 |

TABLE 5
EFFECT OF MP3 COMPRESSION ATTACKS ON BER VALUES

| Bit Rate | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 128 Kbps | 0.4960 | 0.1456 | 0.2392 | 0.3772 |
| 256 Kbps | 0.0076 | 0 | 0 | 0 |
| 320 Kbps | 0 | 0 | 0 | 0 |
| 512 Kbps | 0 | 0 | 0 | 0 |

TABLE 6
EFFECT OF NOISE ATTACKS ON BER VALUES

| Noise | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| -10 dB | 0.5000 | 0.5064 | 0.4876 | 0.4880 |
| -20 dB | 0.4508 | 0.1728 | 0.1764 | 0.3520 |
| -30 dB | 0 | 0 | 0 | 0 |

### 4.7 Cropping

Cropping is the process of cutting the length of a signal with a certain duration located at the beginning or end of the audio file. We analyze the cropping rate to the four types of host files ranging from 70% to 90%. The BER values can be seen in Table 7. Our watermarking system is robust to cropping attack which is shown by the zero BER from the table.

### 4.8 Linear Speed Change (LSC)

LSC is an attack that changes the speed of the audio signal in a linear manner. Changes that occur in the signal depends on the scale of the conversion used, the audio signal that has been attacked can be slow or fast. Table 8 shows the results of our audio watermarking system. From the table, our system is robust to LSC attacks. The test file host3.wav and host4.wav can be said to be robust because BER is very much close to zero. We also conduct subjective tests named Mean Opinion Score (MOS) to determine the quality of watermarked audio by playing the results to 30 correspondents. This test is done by comparing 2 audio files, i.e. host audio and watermarked host audio. It can be seen in Table 9 where our results show that the higher the intensity value the higher the MOS value.

We compare this third scheme with previous schemes proposed in [13]. It can be seen in Table 10 where BER, PSNR, CC, and extracted watermark results according to attack types. Our system is robust against some types of attacks such as LPF, MP3 compression, equalizer and cropping.

TABLE 7
EFFECT OF CROPPING ATTACKS ON BER VALUES

| Cropping | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 70% | 0 | 0 | 0 | 0 |
| 75% | 0 | 0 | 0 | 0 |
| 80% | 0 | 0 | 0 | 0 |
| 85% | 0 | 0 | 0 | 0 |
| 90% | 0 | 0 | 0 | 0 |

TABLE 8
EFFECT OF LSC ATTACKS ON BER VALUES

| LSC Scale | host2.wav | host3.wav | host4.wav | host5.wav |
|---|---|---|---|---|
| 50% | 0 | 0.00004 | 0.00008 | 0 |
| 85% | 0 | 0.00004 | 0.00008 | 0 |
| 90% | 0 | 0.00004 | 0.00008 | 0 |
| 95% | 0 | 0.00004 | 0.00008 | 0 |

TABLE 9
MEAN OPINION SCORE RESULTS

| Tested Files | Intensity 0.1 | Intensity 0.01 | Intensity 0.001 |
|---|---|---|---|
| host2.wav | 4.1818 | 4.2727 | 4.40900 |
| host3.wav | 2.7723 | 4.0909 | 4.31818 |
| host4.wav | 2.9090 | 3.9090 | 4.45454 |
| host5.wav | 3.2272 | 4.0000 | 4.31818 |

## V. CONCLUSION

We analyzed the third scheme of audio watermarking system compared to those in [13]. This paper shows that the systems we designed are robust to some attack types, i.e. LPF 10 kHz, cropping, MP3 compression 256 kHz, and MP3 compression 512 kHz. With no attack type, our system in [13] has BER=0, PSNR=$\infty$ and CC=1. Our third scheme system has BER=0.0012, PSNR=29.2082 and CC=0.9976 for compression and LPF attacks; while with cropping attack this paper shows BER=0, PSNR=$\infty$ and CC=1.

## REFERENCES

[1] R. Petrovic, "Audio signal watermarking based on replica modulation", 5th International Conference TELSIKS'01, 2001.

[2] N. V Lalitha, “DWT-Arnold Transform Based Audio Watermarking,” pp. 196–199, 2013.

[3] A. Al-haj, A. Mohammad, and L. Bata, “DWT – Based Audio Watermarking,” *Int. Arab J. Inf. Technol.*, vol. 8, no. 3, pp. 326–333, 2011.

[4] M. Charfeddine, M. El’Arbi, C. B. Amar, “ A New DCT Audio Watermarking Scheme Based on Preliminary MP3 Study,” International Journal of Multimedia Tools and Applications, 2014.

[5] N.V.Lalitha, P.Vara Prasad, S.Uma Maheshwar Rao, “Performance Analysis of DCT and DWT Audio watermarking based on SVD”, International Conference on Circuit, Power and Computing Technologies [ICCPCT], 2016.

[6] N. V. Lalitha, S. Rao, P. V. Y. J. Sree, “ DWT – Arnold Transform Based Audio Watermarking,” IEEE Asia Pasific Conference on Postgraduate Research in Microelectronics and Electronics (Prime Asia), 2013.

[7] Cai Yong-mei, Guo Wen-qiang, Ding Hai-yang, “An Audio Blind Watermarking Scheme Based on DWT-SVD”, Journal of Software, Vol.8, No.7, July 2013.

[8] R. Alshazly, M. M. Fouad, M. E. Nasr, “ Secure and Robust High Quality DWT Domain Audio Watermarking Algorithm With Binary Image,” International Conference on Computer Engineering & Systems (ICCES), 2012.

[9] I. Safitri, R. R. Ginanjar, A. Rizal, Azizah, “Adaptive Multilevel Wavelet BCH Code Method in the Audio Watermarking System,” International Conference on Controls, Electronics, Renewable Energy and Communications (ICCEREC), 2017.

[10] H. Abdi, “Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD),” Encyclopedia of Measurement and Statistics, 2007.

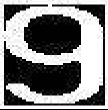
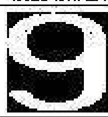
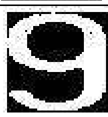
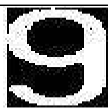
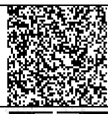
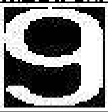
[11] Audio Watermarking Using Arnold Transform, DWT and Modified SVD”, IJIRSET Vol. 5, Issue 7, July 2016.

[12] N.V.Lalitha, Ch.Srinivasa Rao, P.V.Y.JayaSree, “DWT-Arnold Transform Based Audio Watermarking”, IEEE Asia Pasific Confrence on Postgraduate Research in Microelectronics and Electronics, 2013.

[13] Bruguiera A. F. Agradriya, Faisal K. Perdana, Irma Safitri, Ledya Novamizanti, “Audio Watermarking Technique Based on Arnold Transform,” International Conference on Automation, Cognitive Science, Optics, Micro Electro—Mechanical System, and Information Technology (ICACOMIT), 2017.

[14] Min Li, Ting Liang, Yu-Jie He, “Arnold Transform Based Image Scrambling Method”, ICMT, 2013.

TABLE 10
THIRD SCHEME OF AUDIO WATERMARKING TECHNIQUE COMPARED TO [13]

| Attack Type | BER | PSNR | CC | Extracted Watermark |
|---|---|---|---|---|
| No Attack | 0 | ∞ | 1 |  |
| Up sampling 88.2 kHz | 0.5076 | 2.9448 | -0.0131 |  |
| Down sampling 22.05 kHz | 0.5076 | 2.9448 | -0.0112 |  |
| MP3 Compression 256 kHz | 0.0012 | 29.2082 | 0.9976 |  |
| MP3 Compression 512 kHz | 0.0012 | 29.2082 | 0.9976 |  |
| LPF 10 KHz | 0.0012 | 29.2082 | 0.9976 |  |
| Equalizer | 0.4860 | 3.1336 | 0.0273 |  |
| Cropping | 0 | ∞ | 1 |  |